ACS Catalysis

Research Article

pubs.acs.org/acscatalysis

Computer-Aided Engineering of a Transglycosylase for the Glucosylation of an Unnatural Disaccharide of Relevance for Bacterial Antigen Synthesis

Alizée Verges,[†,‡,∥,#] Emmanuelle Cambon,[†,‡,∥,#] Sophie Barbe,[†,‡,∥,#] Stéphane Salamone,[Δ,○,□] Yann Le Guen,[Δ,○,⊥] Claire Moulis,[†,‡,∥] Laurence A. Mulard,[Δ,○] Magali Remaud-Siméon,*[,†,‡,∥] and Isabelle André*[,†,‡,∥]

[†]Université de Toulouse; INSA,UPS,INP; LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France

[‡]CNRS, UMR5504, F-31400 Toulouse, France

[∥]INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, France

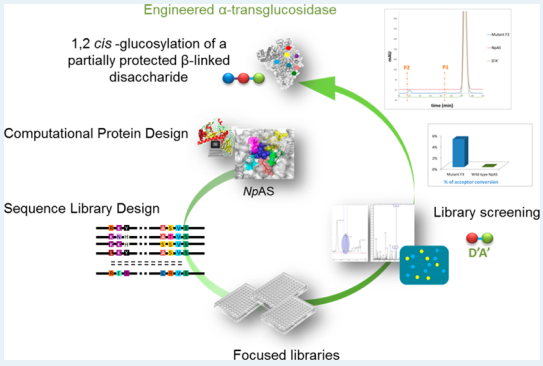
[Δ]Institut Pasteur, Unité de Chimie des Biomolécules, 28 rue du Dr. Roux, F-75724 Paris Cedex 15, France

[○]CNRS UMR3523, Institut Pasteur, F-75015 Paris, France

[⊥]Université Paris Descartes Sorbonne Paris Cité, Institut Pasteur, F-75015 Paris, France

Supporting Information

ABSTRACT: The exploration of chemo-enzymatic routes to complex carbohydrates has been hampered by the lack of appropriate enzymatic tools having the substrate specificity for new reactions. Here, we used a computer-aided design framework to guide the construction of a small, diversity-controlled library of amino acid sequences of an α-transglucosylase, the sugar binding subsites of which were re-engineered to enable the challenging 1,2-*cis*-glucosylation of a partially protected β-linked disaccharide allyl (2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl)-(1→2)-α-L-rhamnopyranoside, a potential intermediate in the synthesis of *Shigella flexneri* cell-surface oligosaccharides. The target disaccharide is not recognized by the parental wild-type enzyme and exhibits a molecular structure very distinct from that of the natural α-(1→4)-linked acceptor. A profound reshaping of the binding pocket had thus to be performed. Following the selection of 23 amino acid positions from the first shell, mutations were sampled using RosettaDesign leading to a subset of 1515 designed sequences, which were further analyzed by determining the amino acid variability among the designed sequences and their conservation in evolutionary-related enzymes. A combinatorial library of 2.7 × 10^4 variants was finally designed, constructed, and screened. One mutant showing the desired and totally new specificity was successfully identified from this first round of screening. Impressively, this mutant contained seven substitutions in the first shell of the active site leading to a drastic reshaping of the catalytic pocket without significantly perturbing the original specificity for sucrose donor substrate. This work illustrates how computer-aided approaches can undoubtedly offer novel opportunities to design tailored carbohydrate-active enzymes of interest for glycochemistry or synthetic glycobiology.



KEYWORDS: *oligosaccharide, amylosucrase, computer-aided library design, enzyme engineering, chemo-enzymatic synthesis, Shigella flexneri*

INTRODUCTION

Enzyme engineering is a major instrument of modern biocatalysis and novel enzyme-based process development.[1,2] Advances in bioinformatics, biophysics, biochemistry, and screening instrumentation provided all their contribution to rationalize or accelerate the optimization, adaptation, and even the conception of enzymes.[3] Learning from the natural evolution process has allowed decrypting some of the natural mechanisms that shape protein fitness landscapes and has served to develop laboratory methods of evolution and to tailor catalysts for synthetic purposes.[4] In particular, computer-aided approaches including sequence alignments combined with or without structural data, phylogenetic analyses, and identification of mutation correlations within protein superfamilies occupy an increasing place in the rationalization of enzyme evolution to design focused libraries, guide amino acid substitutions, and narrow down the size of the library.[5] Also, computational protein design (CPD) is developing quickly,[6−9] and despite existing

Received: August 29, 2014
Revised: January 7, 2015
Published: January 9, 2015

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

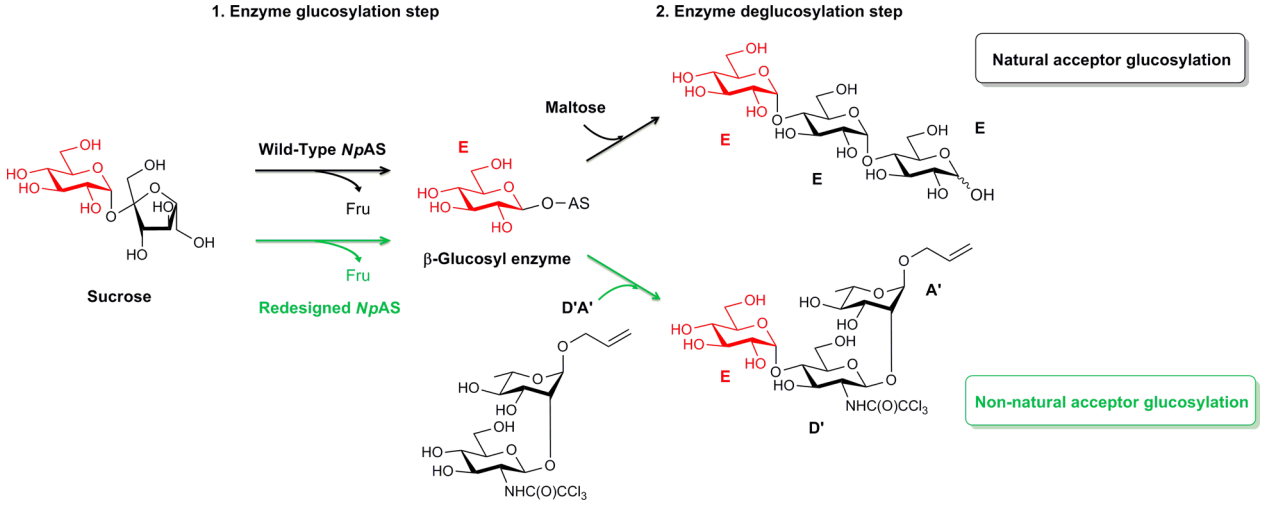


Figure 1. Reaction catalyzed by *Np*AS from sucrose in the presence of natural acceptor (maltose) and targeted reaction by redesigned amylosucrase (in green) using the non-natural acceptor **D′A′**. Catalytic mechanism is decomposed in two sequential steps: the first step (enzyme glucosylation step) leads to the formation of a *β*-glucosyl enzyme intermediate, whereas the second one (enzyme deglucosylation step) corresponds to the transglucosylation of the acceptor molecule (natural or non-natural). *Np*AS: *Neisseria polysaccharea* amylosucrase; sucrose, *α*-D-glucopyranosyl-(1,2)-*β*-D-fructofuranoside; Fru, Fructose; Maltose, Glucopyranosyl-(1−4)-*α*-D-glucopyranose; **A′**, allyl *α*-L-rhamnopyranoside; **D′**, 2-deoxy-2-trichloroacetamido-*β*-D-glucopyranosyl; **E**, *α*-D-Glucopyranosyl

limitations in the treatment of flexibility, sequence-conformation exploration, solvent consideration, and energy functions, it provides extremely useful information to create libraries of defined sequences.[10−17] An increasing number of examples demonstrate their potential to draw evolutionary paths leading to desired catalytic functions with a good efficacy and accuracy. However, these computer-aided approaches have been scarcely applied to carbohydrate-active enzymes (CAZymes), although they could significantly accelerate the exploration of the combinatorial mutations, especially when the engineering of multiple carbohydrate binding sites is necessary to confer the desired substrate specificity to the enzyme.[18]

Although the number of natural CAZymes able to act in a regio- or stereospecific manner on a vast range of glycosidic molecules is continuously expanding in gene databases, access to catalysts that will address a particular need is not a trivial issue. In particular, the lack of appropriate enzymatic tools with requisite substrate and bond specificities has hampered extensive exploration of enzymatic or chemo-enzymatic routes to complex carbohydrates and the design of biocatalysts, which in combination with chemistry could help bypassing some constraints encountered in glycochemistry, such as the need of protection of highly reactive hydroxyl groups. Such constraints can be integrated into the (re)design of enzymes, purposely "tailored" to enter the chemo-enzymatic pathway at a specific and programmed stage and thus enable thorough exploration of novel pathways toward defined carbohydrate-based structures.[19,20]

This is the challenge addressed in this study where a computer-aided approach was followed to redesign several sugar binding subsites of a sucrose-utilizing transglucosylase, the amylosucrase from *Neisseria polysaccharea* (*Np*AS). The objective was to adapt the enzyme to the transglucosylation of a non-natural disaccharide, structurally highly divergent from its natural substrates, with the long-term goal of developing in vitro chemo-enzymatic pathways to access complex oligosaccharides featuring *α*-D-glucopyranosyl side chains. Indeed, *α*-D-glucosylation is frequently encountered in natural polysaccharides either in the backbone or as side-chain substitutions. The latter situation is nicely exemplified in *Shigella flexneri*, a family of Gram negative bacteria. Most of the O-specific polysaccharides (O-SP) of the different *S. flexneri* types and subtypes show repeating units with the same linear tetrasaccharide backbone, and differing mainly in the site-specific *α*-D-glucosylation and O-acetylation of this common tetrasaccharide.[21,22] Recently, we demonstrated that the active site of the *α*-retaining *Np*AS could be reshaped to enable the site-selective *α*-D-glucosylation of two non-natural acceptor monosaccharides, namely, methyl *α*-L-rhamnopyranoside and allyl 2-acetamido-2-deoxy-*α*-D-glucopyranoside.[19,20] A more ambitious design relying on the re-engineering of several binding subsites of *Np*AS was undertaken here to render the enzyme able to glucosylate disaccharide **D′A′**, featuring a partially protected D-glucosamine (**D′**) *β*-(1,2)-linked to a partially protected L-rhamnoside (**A′**) (Figure 1, Figure S1 of the Supporting Information). The unnatural disaccharide acceptor (**D′A′**) was designed so as to open the way to a set of serotype-specific *S. flexneri* O-SP motifs after appropriate enzymatic *α*-D-glucosylation (Figure S1 of the Supporting Information). In the following, emphasis was put on the conversion of the **D′A′** acceptor into the **ED′A′** trisaccharide, featuring an *α*-D-glucopyranosyl residue (**E**) at O-$4_{D'}$ (Figure 1). This glucosylation pattern is present in the O-SP from *S. flexneri* serotypes 1a, 1b, and 1d.[21,22] The CPD approach was first used for the generation of a primary sequence library, which was then rationally reduced in size by integration of additional information regarding conserved amino acids or correlated pairs of amino-acids found in homologous enzymes. Reasons for adopting such a tactic were motivated by the objectives addressed in this work: (1) conserving the original specificity toward donor substrate, while (2) creating a novel specificity for a complex acceptor substrate not recognized by the parental wild-type enzyme; (3) requiring thus the redesign at multiple binding subsites of a very constrained active site; and (4) needing of downsizing mutant libraries due to the availability of only low- and medium-throughput screening assays.

Using the framework disclosed herein, a library of defined amino-acid sequences encoding only for the desired combinations was constructed and screened. An *α*-transglycosylase,

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

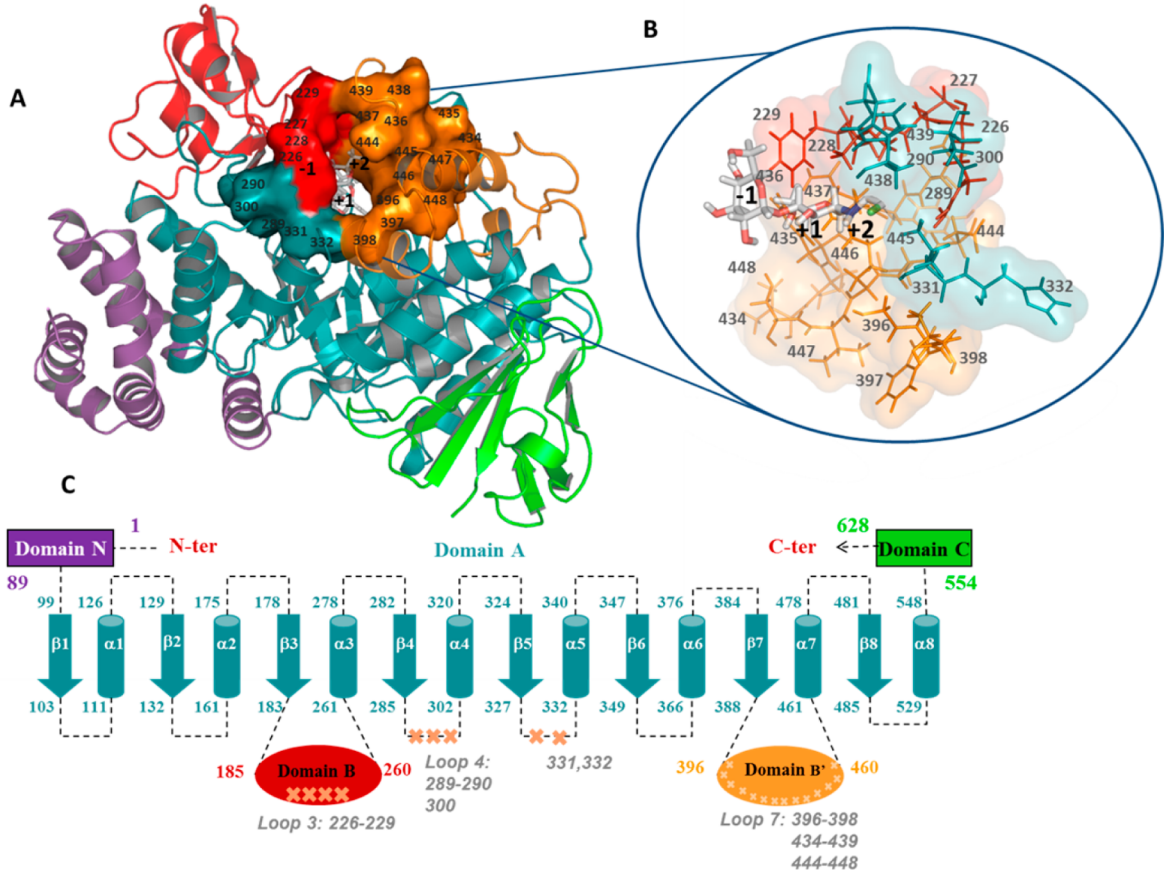


Figure 2. View of amylosucrase organization. Constitutive domains of *Np*AS are colored respectively in purple for the N-domain, in cyan for the A-domain ((α/β) catalytic barrel), in red for the B-domain (loop 3 of catalytic barrel), in orange for the B′ domain (extension of the loop 7 from catalytic barrel), and green for the C-domain. (A) Overall view of *Np*AS in complex with docked **ED′A′** showing the 23 selected positions forming the binding site; (B) cross-sectional view of the active site with **ED′A′** occupying subsites −1, +1 and +2; (c) schematic representation of the amylosucrase topology and location of the 23 selected position.

whose natural specificity is the synthesis of α-(1,4)-linked glucopyranosyl polymers from sole sucrose,[23−25] was converted into an enzyme displaying a novel substrate specificity for the α-D-glucosylation at OH-$4_{D'}$ of the β-(1,2)-linked disaccharide **D′A′** for which there is no equivalent reaction yet reported in the literature.

RESULTS AND DISCUSSION

Candidate Enzyme and Selected Disaccharide. Adding to the fact that a wealth of structural information is already available on *Np*AS enabling the use of computer-aided engineering strategies, this enzyme appeared to be a promising and malleable candidate for the challenge undertaken herein, which is the conception of a mutant able to regiospecifically glucosylate, from sucrose donor, a lightly protected disaccharide acceptor. As seen in Figure 1, the structure of disaccharide **D′A′** is very different from that of the natural acceptor (maltose). It is noteworthy that this molecule is not recognized at all by the wild-type enzyme, and its accommodation necessitates major redesign of both +1 and +2 subsites to allow its glucosylation during the second step of the catalytic reaction (Figure 1). In contrast to most chemo-enzymatic syntheses of complex oligosaccharides described to date,[26−29] the strategy under investigation is based on an early stage enzymatic glucosylation, followed by chemical chain elongation of the obtained product. Toward this aim, it is critical that the products of the enzymatic glucosylation step are compatible with further chemical transformation into building blocks permitting chain elongation at both the reducing and nonreducing ends. We reasoned that the use of lightly protected acceptor substrates should in theory fulfill these criteria without preventing enzyme remodeling. In the following, disaccharide **D′A′**, namely, allyl (2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl)-(1→2)-α-L-rhamnopyranoside (Figure 1), corresponding to a protected version of part of the backbone tetrasaccharide repeating unit of several *S. flexneri* surface polysaccharides (Figure S1 of the Supporting Information),[21,22] was designed as the selected acceptor. Besides being orthogonal to most protecting groups, the allyl moiety does not bring in major steric hindrance. Moreover, it was previously shown to be compatible with *Np*AS remodeling and successful chain elongation of the disaccharide resulting from the enzymatic glucosylation step of a monosaccharide acceptor.[19] The remote effect of acetamides on chemical glycosylation reactions was reported,[30] and the synthesis of oligosaccharides bearing *N*-acetylated aminosugars usually involves building blocks whereby the acetamido function is masked. In this context, we have successfully used the trichloroacetyl moiety as a N-protecting group for the aminosugar building blocks involved in the chemical synthesis of oligosaccharides representative of the O-SP from different *S. flexneri* types and subtypes.[31,32] In agreement with our previous achievements, the 2_D acetamide in disaccharide **DA** is masked as a trichloroacetamide moiety in acceptor **D′A′**.

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

Design and Construction of Directed Sequence Libraries. Crystallographic structure of *Np*AS in complex with the maltoheptaose product[25] was used as a template to manually force docking of the desired glucosylated product (allyl α-D-glucopyranosyl-(1→4)-(2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl)-(1→2)-α-L-rhamnopyranoside, **ED′A′**) into the enzyme active site. Examination of docked **ED′A′** showed that all binding interactions with the glucopyranosyl moiety (**E**) at the −1 subsite, which is responsible for enzyme specificity toward sucrose, were preserved in comparison with crystallographic data on complexes. After excluding all essential amino acid residues for either catalysis or specific recognition of sucrose, we selected altogether 23 mutable positions from the first shell of amino acid residues of subsites +1, +2, and +3, which were likely by mutation to favor recognition of **D′A′**. These 23 positions are located in structurally distinct regions of the enzyme, mainly on loops 3, 4, and 7, which surround the catalytic center (Figure 2). If all 23 positions were considered to be mutable by one of the 20 possible amino acid residues, the theoretical space would be as large as 8.4×10^{29}, clearly out of reach of available screening methods. Prior methods focusing on the use of degenerate oligonucleotides encoding for a limited set of amino acids (such as NDT or NDK codons encoding for only 12 amino acids) have been proposed to reduce the library size.[33] However, these methods randomly introduce a reduced set of amino acids without taking into account the structural context and the mechanistic constraints. To tackle the customized construction of a small, semi-random, and diversity-controlled library, we proposed to adopt an approach enabling to focus only on the most promising mutations derived from computational predictions and that were refined by additional integration of structural and biochemical knowledge.[23−25,34] Considering the 23 selected positions and using the enzyme design protocol described in Figure 3, we sampled mutations at each position by performing 20 000 independent runs of RosettaDesign. The 20 000 generated sequences were subsequently scored and filtered using various criteria including predicted favorable enzyme-ligand binding interaction and complex stability. This filtering process yielded a subset of 1515 designed sequences. At this stage, we decided to not directly generate this library of mutants. Indeed, in some cases, the predicted structures differed only by very small scoring differences that might have been irrelevant given all approximations made in the calculations. On the other hand, the combinatorial space of this library remains huge, around 5.5×10^{11} sequences and still too broad to be handled experimentally. The reduction of library size was thus pursued on the basis of a specific sequence analysis consisting in determining the amino acid variability within the 1515 designed sequences and their conservation in the evolutionary related enzymes to subfamily GH13_4 of glycoside-hydrolases,[34] which share a common specificity toward sucrose substrate donor.

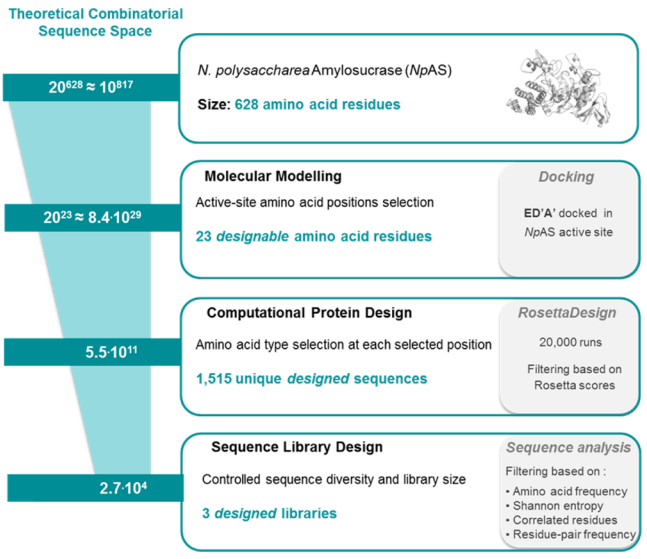


Figure 3. Computer-aided approach for enzyme and sequence library design. The size of the theoretical combinatorial sequence space is indicated at each stage of the strategy.

On the basis of the amino acid occurrence observed for each position in the 1515 designs, mutations that had low occurrence (≤10%) were discarded. Whenever the presence of wild-type amino acids was found over or equal to 5%, it was maintained at the given position in the final designed library. Although at some positions the wild-type amino acid was not (or only scarcely) observed in the design results (frequency of occurrence less than 5%), it was kept or reintroduced in the final library at this position when it was found highly conserved in the GH13_4 subfamily (Shannon entropy less than 0.2%). The decision to only reintroduce wild-type amino acids at very specific positions is a compromise considering that a high occurrence in the subfamily GH13_4 may be critical for sucrose substrate specificity but that the preservation of wild-type signature at all the positions would have led to a too large library. Indeed, if the wild-type amino acid had been conserved at each designable position, the size of the library would have been as large as ca. 10^7 compared to 1.2×10^5 using a position-dependent reintroduction of wild-type amino acids. To further reduce the size of the combinatorial library, pairwise correlation analysis of predicted mutations was carried out. On this basis, only significant interposition correlations (residue-pair frequency ≥ 5%) were kept in order to only promote beneficial combinations of mutations. Such removal of correlated pair combinations displaying low residue-pair frequency allowed narrowing down further the combinatorial space size to 2.7×10^4 sequences. The expected outcome of this approach was to achieve a sharper exploration of the diversity in the regions of higher interest in comparison to the exorbitant size of theoretical combinatorial libraries generated if all 20 amino acid mutations had been allowed at each of the 23 positions (corresponding in theory to 20^{23} possible sequences).

To do so, diversity of the library was controlled in the library construction by mixing synthetic oligonucleotides that encoded for most frequent amino acid mutations and taking into account favorable inter-residue correlations observed in sequences predicted by computational design. Considering the 23 selected positions from the first shell and using a computational approach involving the different filters above-described (Figure 3), the diversity was drastically narrowed down from 20^{23} theoretical possible sequences to 2.7×10^4. To ensure the exploration of all the designed sequences, three sublibraries were constructed (Table 1).

In the designed libraries, the amino acid sequences were expected to contain from 9 up to 17 mutations, avoiding any representation of the parental gene. To construct the libraries, the gene was first amplified using six couples of nonmutated primers positioned all along the sequence to generate six DNA wild-type fragments surrounding the five regions containing designed mutations (Step 1 in Figure S2 of the Supporting Information). These fragments were then purified and used as primers in subsequent PCR reactions in which both wild-type

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

Table 1. Summary of the Predicted Mutations and Authorized Combinations of Mutations at the 23 Selected Positions in the Construction of Sequence Libraries 1−3

Targeted positions	226	227	228	229	289	290	300	331	332	396	397	398	434	435	436	437	438	439	444	445	446	447	448
Wild-type amino acid	R	E	I	F	A	F	E	V	H	G	W	T	V	P	F	Q	Y	N	D	C	R	V	S
Authorized amino acid Library 1	K or L	K or E	H or Y	D or A or F	S	H	I or V	G or T	H	G or S	W	V	R or V	P	D or S or F	R or S	Y	D	D or E	A	A or R or S	V	S
Authorized amino acid Library 2	K or L	K or E	V or Y	D or A or F	A	Y	V	G or T	H	G or S	W	V	R or V	P	D or S or F	R or S	Y	D	D or E	A	A or R or S	V	S
Authorized amino acid Library 3	K or L	K or E	V or Y	D or A or F	I	Y	I	G or T	H	G or S	W	V	R or V	P	D or S or F	R or S	Y	D	D or E	A	A or R or S	V	S
Targeted structural domain	Domain B (Loop 3)						Domain A (Loop 4)			Domain B' (Loop 7)													

and mutated primers were used to amplify longer fragments incorporating the mutations (Step 2 in Figure S2 of the Supporting Information) with an equimolar representation of each desired mutation. The amplicons obtained from step 2 were mixed two by two (Step 3 in Figure S2 of the Supporting Information). At step 4, the five amplified long fragments incorporating the mutations were mixed together in equimolar amounts. The entire gene was finally amplified.

Primary Screening of the Library for Utilization of Sucrose As Donor Substrate. *Screening Results.* The libraries were first screened on sucrose using a colorimetric pH-based assay developed earlier on solid medium,[19,20,35] which enables the detection of the release of acids upon fructose uptake and decreases the number of mutants to be subsequently screened for **D′A′** glucosylation. Out of the 63 000 screened clones, the number of active clones represented only 0.09% of the overall library, showing the drastic impact of some mutations on proper catalytic recognition of sucrose or enzyme viability. Altogether, 59 sucrose-active variants were identified. Of these, 55 variants were confirmed by DiNitroSalicylic acid (DNS) assays (Table 2).

Analysis of these 55 variants revealed that about 25% and 70% of active clones from library 2 and 3 are more active on sucrose than the wild-type enzyme (Figure S3 of the Supporting Information). The only mutant identified from library 1 is less active than wild-type *Np*AS. Overall, 28 clones originating mostly from library 3 were found with equal or higher activity on sucrose cleavage (see Figure S3 of the Supporting Information).

To discriminate transglucosylases from sucrose-hydrolases, the reaction products were analyzed by HPLC. Overall, 25 of the 55 variants (45.5%) synthesized sucrose isomers and malto-oligosaccharides. Thirteen mutants synthesized products differing from those produced by parental wild-type *Np*AS. Seventeen mutants (30.9%) behaved as pure sucrose-hydrolases. The thermal stability of active mutants purified in deep-well plates was assessed by differential scanning fluorimetry (DSF) in the absence of a bound ligand using the procedure previously reported[20] (Figure S4 of the Supporting Information). Overall, T_m values of the 55 mutants ranged from 43.7 ± 0.2 °C to 51.5 ± 0.1 °C. Notably, most mutants resulting from library 2 and the mutant from library 1 turned out to be more stable than wild-type *Np*AS (Tm of 48.6 °C), although overall, these mutants were less active on sucrose. Conversely, mutants from library 3 had lower T_m values than wild-type *Np*AS but were, however, more active on sucrose. Altogether these results seem to indicate that gain in mutant stability is often associated with a loss of activity on sucrose as previously observed for F290K, A289P-F290L, A289P-F290C, and A289P-F290I mutants of *Np*AS.[19,20]

Table 2. Experimental Screening of Libraries 1, 2, and 3 on Sucrose and Analysis of Active Clones

	library		
	1	2	3
theoretical number of clones	13 824	6912	6912
number of screened clones	~23 000	~23 000	~17 000
number of active clones on sucrose using pH-based solid medium assay	5	24	30
number of active clones on sucrose using the DNS assay	1	24	30
% of overall hit rate	4.3×10^{-3}	1.0×10^{-3}	1.7×10^{-3}
average number of mutations in active clones	13[a]	7.8	9.5
minimum number of mutations in active clones	-	5	5
maximum number of mutations in active clones	-	13	17

[a]One single mutant was isolated from library 1.

Conformity of Constructed Libraries with Computational Predictions. The 55 selected clones were sequenced (Table S2 of the Supporting Information). On average, clones were found to contain 8.8 mutations per clone, with each clone displaying a different mutational rate. Among all libraries, library 3 yielded a higher number of amino acid mutations per clone (Table 2). In all, 20% and 16.7% of sequences from libraries 2 and 3, respectively, are totally conformed to the design. The maximum number of errors in the designed sequences is 5 and they were introduced most of the time at the same positions (331, 398, 437, 439 or 445). About 73% and 76% of the encoded amino acid sequences from libraries 2 and 3, respectively, contained at the

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

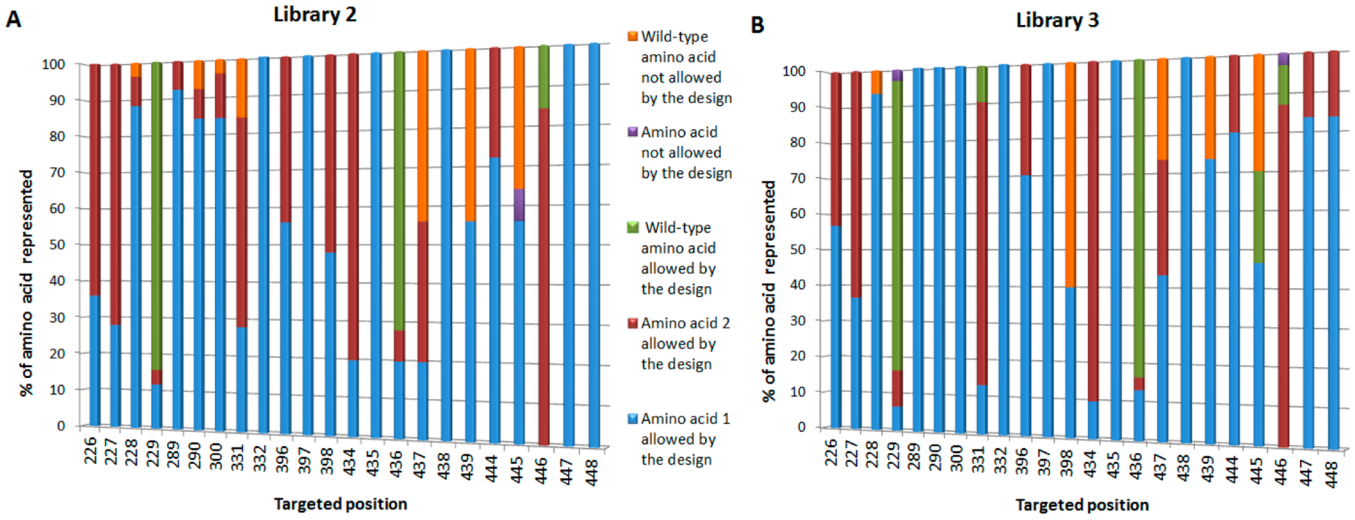


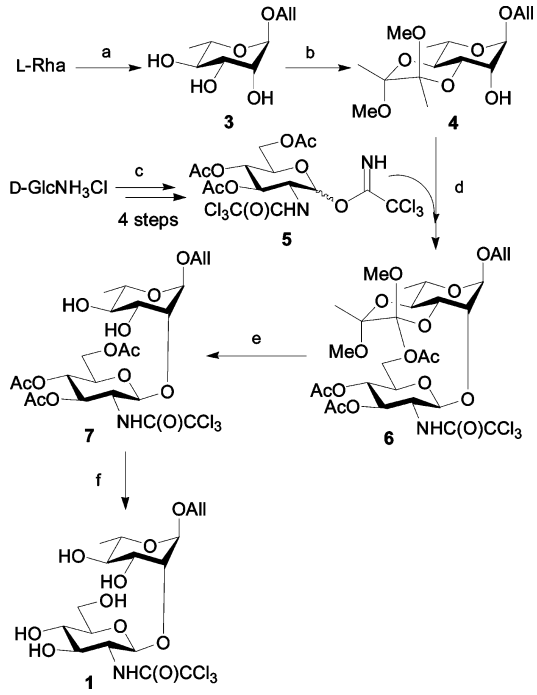
Figure 4. Amino acid composition by position of sucrose-active hits resulting from screening of designed libraries: library 2 (A), library 3 (B).

most three errors compared to the design. The variation at each position is reported in Figure 4. Amino acid mismatches correspond to wild-type amino acid incorporations or substitutions not allowed by the design. Indeed, several sequences from libraries 2 and 3 contained undesired wild-type amino acids at positions 228, 290, 300, 331, 437, 439, and 445 with a particularly higher frequency at positions 437, 439, and 445. These discrepancies may be attributed to the length of the oligonucleotide primers that could affect the proper hybridation of the fragments (case of the region comprised between positions 434 and 448). A higher error rate is also observed when mutations are located at primer extremities (case of mismatching found at position 398, library 3) (Figure 4).

Site-Selective α-D-Glucosylation of Disaccharide D′A′ (1). *Disaccharide Synthesis.* Multigram amounts of disaccharide **D′A′** (**1**) were necessary to allow for library screening and active mutant analysis. Toward this aim, disaccharide **1** was chemically synthesized in 9 steps starting from commercially available L-rhamnose and D-glucosamine (Scheme 1). The use of ester and acetal protecting groups was favored so as to prevent any interference with the allyl aglycon and trichloroacetamide moiety at the final deprotection step. Thus, the known allyl rhamnoside[36] **3** was protected regioselectively at position O-3 and O-4 by taking advantage of the selectivity of the 1,2-diacetal chemistry for 1,2-diequatorial diol systems.[37,38] The preliminary conversion of butane-2,3-dione into the more reactive species 2,2,3,3-tetramethoxybutane[39] was found advantageous, and alcohol **4** was isolated in a good 88% yield upon reaction of triol **3** with the latter. TMSOTf-mediated glycosylation of acceptor **4** with the known trichloroacetimidate glucosaminyl donor[40] **5** yielded the crystallized disaccharide **6** in a non-optimized 83% yield, when run on a multigram scale. Acidic hydrolysis of the acetal protecting group gave diol **7** (86%), and subsequent deacetylation of the latter furnished the target allyl glycoside **1** (92%).

Secondary Screening of Mutants on D′A′ Glucosylation and Identification of Hits. All 55 sucrose-active mutants and wild-type *Np*AS were tested for their ability to glucosylate disaccharide **D′A′**. One mutant (F3) glucosylated **D′A′** with a conversion of 2%. LC-MS analysis of the enzymatic mixtures indicated a molecular weight increase by 162 g.mol^{-1} (**P1**) and 324 g.mol^{-1} (**P2**) for the novel UV-detectable compounds (**P1**: 670 and **P2**: 832 g.mol^{-1}, respectively), corresponding to the transfer of one or two glucosyl units onto **D′A′** (Figure 5).

Scheme 1. Chemical Synthesis of Disaccharide D′A′ (1) from L-Rhamnose and D-Glucosamine Hydrochloride[a]



[a]Reagents and conditions: (a) see ref 36; (b) 2,2,3,3-Tetramethoxybutane, CSA, MeCN, reflux, 30 min (88%); (c) see ref 40; (d) TMSOTf, DCM, −15 °C → rt, 40 min (83%); (e) 90% aq TFA, rt, 50 min (86%); (f) MeONa, MeOH, rt, 2 h (92%) ; CSA: Camphorsulfonic acid, TMSOTf: trimethylsilyltrifluoromethanesulfonate, DCM: dichloromethane, rt: room temperature; TFA: trifluoroacetic acid.

Characterization of the Mutant Displaying a Novel D′A′ Specificity and Synthesized Products. Sequencing of the mutant F3 that displays a totally novel substrate specificity toward **D′A′**, not detected for the wild-type *Np*AS, revealed the presence of 7 mutations (R226L-I228V-F290Y-E300V-V331T-G396S-T398V). This mutant has a T_m of 45.5 °C instead of 48.6 °C for the wild-type enzyme. Reactions in the presence of sole

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

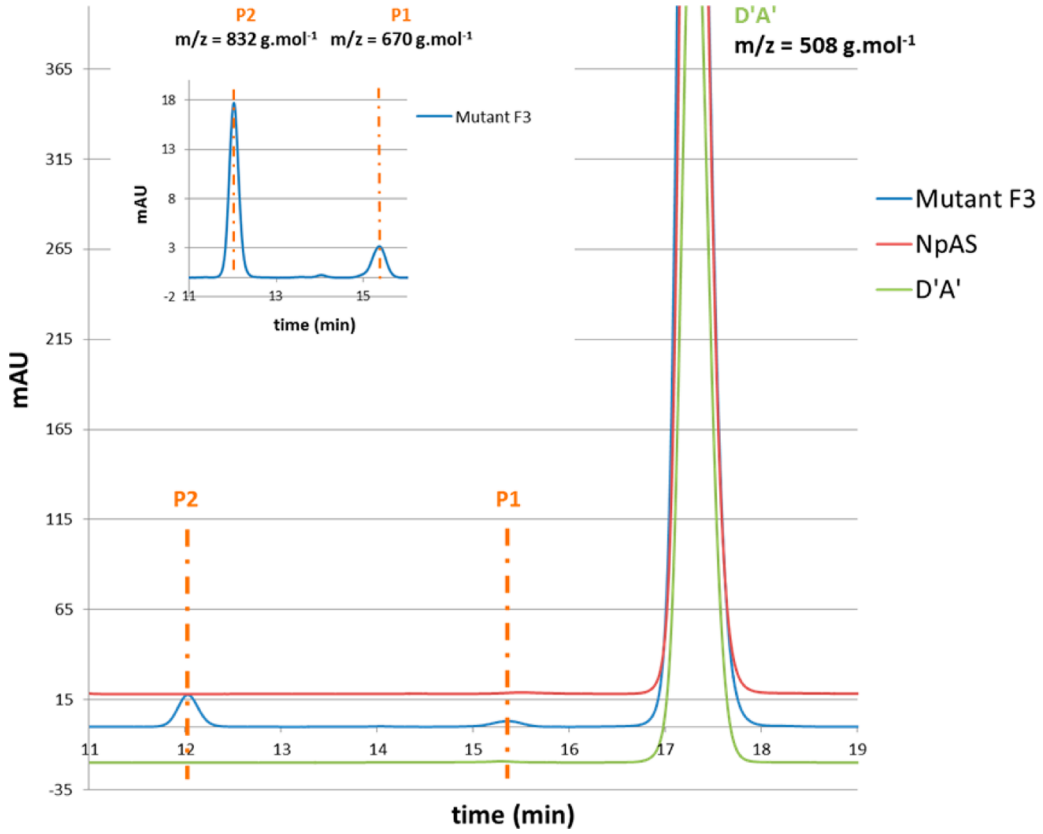


Figure 5. Comparison of product profiles obtained with wild-type *Np*As (red) and mutant F3 (blue) after 24 h using 146 mM sucrose and 146 mM **D′A′**. HPLC analysis was performed (C18 phase column (Venusil AQ 50 mm) with acetonitrile/water (17:83, v/v) as eluent, 30 °C; 1 mL/min) with UV detection at 205 and 220 nm. The HPLC profile of **D′A′** (t_r = 17.3 min) is shown in green for reference. **P1** (t_r = 15.3 min) and **P2** (t_r = 12.1 min): glucosylation reaction products of 670 and 832 g·mol^{-1}, respectively. A zoom of the product **P1** and **P2** is embedded in the figure.

sucrose showed that the specific activity on sucrose of the mutant was 1.8-fold higher than that of the wild-type enzyme. As seen in Figure 6A, the mutant F3 produces a higher amount of turanose (34.2% versus 25.3% for wild-type), more maltose (14.9% vs 6.1% for wild-type) and also shorter oligosaccharides from DP 3 to 7 (29.6% vs 3.1% for wild-type) compared to the wild-type enzyme. Indeed, no insoluble α-glucans (DP > 25, 20%) were produced. When **D′A′** (146 mM) was added to the reaction medium, the product profile was modified due to the glycosylation of **D′A′** leading to products **P1** and **P2**, representing 0.6% and 3.7% of the transferred glucopyranosyl residues, respectively. Glucosylation of **D′A′** was also accompanied by an increase of small size maltooligosaccharides and trehalulose production and a pronounced decrease of turanose synthesis (twice less than that of the wild-type) (Figure 6B).

Following mass spectrometry analysis, **P1** and **P2** molecular structures were determined by NMR analysis of tetrasaccharide **P2**. Data showed that first glucosylation of **D′A′** occurred at OH-$4_{D'}$ to provide allyl α-D- glucopyranosyl-(1→4)-2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl-(1→2)-α-L-rhamnopyranoside, also named **ED′A′** (or **P1** in HPLC analysis). Subsequent chain elongation of the product of glucosylation **ED′A′** at OH-6_E provided allyl α-D-glucopyranosyl-(1→6)-α-D-glucopyranosyl-(1→4)-(2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl)-(1→2)-α-L-rhamnopyranoside, also named **EED′A′** (or **P2** in HPLC analysis).

The production of **ED′A′** and **EED′A′** using various sucrose/**D′A′** ratio was investigated (Table 3). Increasing amounts of **D′A′** allows to produce the highest amount of glycosylated

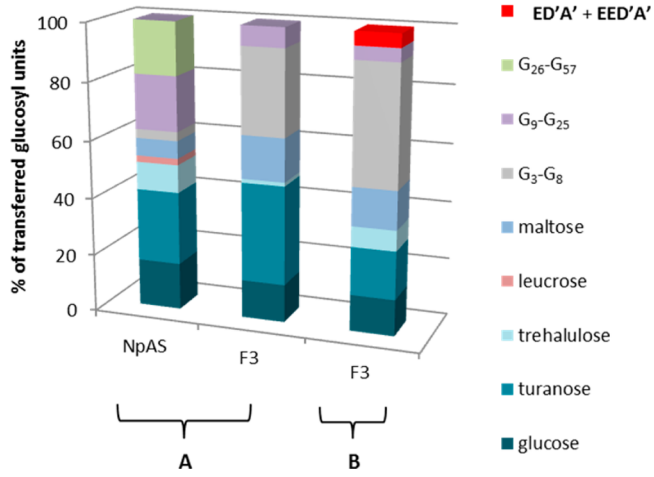


Figure 6. Yields of reaction products analyzed by HPAEC-PAD (Dionex Carbo-Pack PA100 column; 30 °C; 1 mL/min; gradient of sodium acetate (from 6 to 300 mM in 28 min) in 150 mM NaOH) of wild-type *Np*AS and mutant F3 after 24 h reaction. (A) In the presence of sole sucrose (146 mM) or (B) in the presence of both sucrose (146 mM) and **D′A′** (146 mM). G: Glucose; **ED′A′**: allyl α-D-glucopyranosyl-(1→4)-(2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl)-(1→2)-α-L-rhamnopyranoside; **EED′A′**: allyl α-D-glucopyranosyl-(1→6)-α-D-glucopyranosyl-(1→4)-(2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl)-(1→2)-α-L-rhamnopyranoside.

acceptors (2.2 mM of **ED′A′**, 8.9 mM of **EED′A′** at sucrose/acceptor ratio 1/8) in 24 h. The best compromise between the

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

Table 3. Yields of D′A′ Conversion and Quantification of Mono-Glucosylated (P1) and Di-Glucosylated (P2) Products by the Mutant F3 Using Different Donor/Acceptor Molar Ratios after 24 h of Reaction

	sucrose/**D′A′** ratio ([suc] = 73 mM)				sucrose/**D′A′** ratio ([suc] = 146 mM)	sucrose/**D′A′** ratio ([**D′A′**] = 73 mM)		
	1/1	1/2	1/4	1/8	1/1	2/1	4/1	8/1
D′A′ conversion (%)[a]	2	2.17	1.7	1.9	3.0	3.0	4.3	5
monoglucosylated **D′A′** (**ED′A′ or P1**)[b], mM	0.3	0.5	1.1	2.2	0.6	0.3	0.3	0.3
diglucosylated **D′A′**(**EED′A′ or P2**)[c], mM	1.2	2.5	3.7	8.9	3.7	1.9	2.8	3.3

[a]Conversion of **D′A′** was calculated as follows: [P1] mM + [P2] mM/[**D′A′**] mM *100 [b]Monoglucosylated **D′A′** (**ED′A′**, **P1**) determined from HPLC analysis (C18 phase column (Venusil AQ 50 mm) with acetonitrile/water (17:83, v/v)) using calibration curves established with **D′A′** [c]Diglucosylated **D′A′** (**EED′A′**, **P2**) determined from HPLC analysis (C18 phase column (Venusil AQ 50 mm) with acetonitrile/water (17:83, v/v)) using calibration curves established with **D′A′**

conversion rate and yield of acceptor reaction products is obtained for the sucrose/acceptor ratio of 1/1, starting with 146 mM sucrose. Regardless of the sucrose/acceptor ratio used, the production level of the monoglucosylated acceptor (up to 2.2 mM) remained always lower than that of **EED′A′**, suggesting that **ED′A′** is a better acceptor than disaccharide **D′A′**.

The steady-state kinetic parameters (k_{cat}, K_m, k_{cat}/K_m) of transglucosylation reactions of both the wild-type *Np*AS and the mutant F3 were first determined upon varying sucrose concentration from 5 to 600 mM (Table 4). The mutant F3

Table 4. Apparent Kinetic Parameters Determination of Mutant F3 for Sucrose Donor only (from 5 to 600 mM) and in the Presence of Both Sucrose and D′A′ (See Foonote[a])

sucrose only (from 5 to 600 mM)	*Np*AS	mutant F3
K_m apparent[b] (mM)	50.2	21.58
k_{cat}[b] (s^{-1})	1.3	0.82
k_{cat}/K_m[b] (s^{-1} mM^{-1})	0.026	0.040
sucrose (from 5 to 600 mM) + **D′A′** (200 mM)		
$K_{m,\,suc}$[c] (mM)	–[e]	71
$k_{cat,\,suc}$[c] (s^{-1})	–[e]	1.75
$k_{cat}/K_{m,\,suc}$[c] (s^{-1} mM^{-1})	–[e]	0.024
D′A′ (from 5 to 200 mM) + sucrose (200 mM)		
$k_{cat}/K_{m,\,D'A'}$[d] (s^{-1} mM^{-1})	–[e]	3.5×10^{-4}

[a]Sucrose concentration at 200 mM and various **D′A′** concentrations from 5 to 200 mM or 200 mM **D′A′** and sucrose from 5 to 600 mM. [b]Catalytic parameters (K_m, k_{cat}, k_{cat}/K_m) were determined in the presence of sucrose only from 5 to 600 mM. [c]Catalytic parameters (K_m, k_{cat}, $k_{cat}/K_{m,\,suc}$) were determined by measuring the initial rate of fructose production in the presence of constant acceptor concentration (200 mM) and sucrose from 5 to 600 mM. [d]Catalytic efficiency $k_{cat}/K_{m,\,D'A'}$ was determined by measuring the initial rate of **ED′A′** formation from 200 mM sucrose (fixed concentration) and 5 to 200 mM **D′A′**. As **D′A′** was not soluble at concentrations higher than 200 mM, no saturation could be observed and only $k_{cat}/K_{m,\,D'A'}$ was determined from the slope of the linear regression. [e]Not determined due to **D′A′** non recognition.

showed a standard saturation kinetic behavior like the wild-type enzyme. The K_m value was found diminished by 2.3 fold in comparison to the *Np*AS value, indicating a higher affinity of the mutant for sucrose. Conversely, the k_{cat} value of the mutant F3 is decreased by 1.6 fold compared to parental *Np*AS. The catalytic efficiency of the enzymes (k_{cat}/K_m values) was determined by linear regression of the transglucosylation initial rates of measured as variable substrate concentration (sucrose donor or acceptor) (Table 3). In the presence of 200 mM **D′A′** acceptor (considered as a saturating concentration), the catalytic efficiency for sucrose consumption decreased and this appeared to be mainly due to the increase of K_m, suggesting that the presence of disaccharide **D′A′** affects sucrose binding.

Kinetic parameters were also determined in the presence of sucrose (200 mM) and **D′A′** acceptor (concentrations varying from 5 to 200 mM). Due to the limited solubility of **D′A′**, the enzyme could never be saturated, indicating a very low affinity of **D′A′** for mutant F3 leading to a poor catalytic efficiency value.

Although the catalytic efficiency of the mutant is low, we can expect that additional rounds of directed evolution could help to further improve it. Nevertheless, the design, based on molecular modeling predictions, successfully conferred the novel and desired acceptor specificity without substantially affecting the original activity toward sucrose.

Structural Insight on Novel Substrate Specificity. Our efforts to crystallize the mutant F3 were unsuccessful. Therefore, a three-dimensional model of the mutant was first built using the X-ray structure of the parental *Np*AS as template.[23,25,41] Both sucrose and trisaccharide **ED′A′**, the reaction product of interest, were subsequently docked into the active site. Inspection of the models showed that nearly all seven mutations were carried by loops 3 and 4 which have been shown to be highly flexible by molecular dynamics simulations[19,20] and assumed to play a determinant role in the catalytic mechanism of the enzyme. As a result of the introduction of these mutations, the active site of the enzyme is widened up at subsites +1 and +2 enabling the accommodation of the **ED′A′**, and in particular its *N*-trichloroacetyl and allyl substituents, while the productive binding of sucrose is not disturbed, as confirmed by experiment (Figure 7). Indeed, detailed analysis of interactions established between the mutant F3 and **ED′A′** revealed that mutation of Glu300 and Arg226, which are involved in a salt bridge interaction in the parental wild-type enzyme, are mutated by smaller aliphatic residues, valine and leucine, respectively, likely enhancing the plasticity of the active site and favoring the productive binding of **D′A′**. Mutation of Thr398 by a valine residue enables the nesting of the allyl aglycon in a small depression, while mutation of Gly396 by a serine residue brings an additional hydrogen bonding interaction which stabilizes further the allyl group. The mutation V331T also enables to introduce a stabilizing hydrogen bonding interaction with the $2_{D'}$ *N*-trichloroacetyl group. Introduction of the Tyr residue at position 290 requires the creation of enough space, provided by the I228V mutation, to accommodate the extra hydroxyl group compared to the Phe residue present in wild-type enzyme. Altogether these seven mutations enabled a subtle interplay between the structure, the flexibility, the stabilization and the molecular recognition, which cannot yet be predicted altogether by computational methods. As a result, the combination of these mutations enabled the reshaping of the enzyme active site that

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

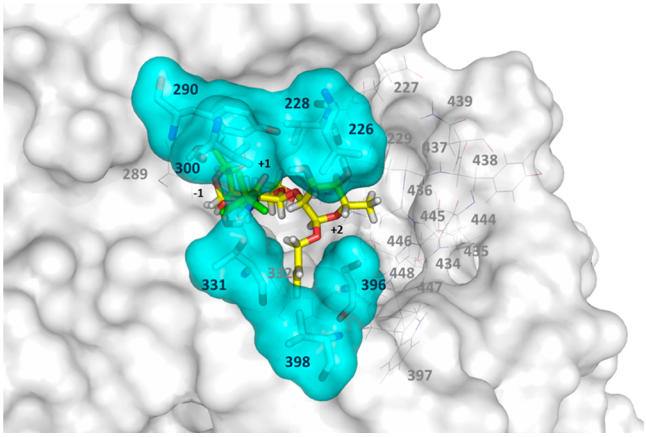


Figure 7. Molecular docking of **ED′A′** in the active site of a built 3D model of the mutant F3 displaying novel acceptor specificity toward **D′A′**. The 7 mutations (R226L-I228V-F290Y-E300V-V331T-G396S-T398V) are highlighted in cyan color (hydrogen atoms have been omitted for clarity purpose). In gray lines are shown the remaining positions targeted by the computational design but not mutated in the mutant F3.

was beneficial for the glucosylation of **D′A′** while preserving sucrose recognition. Interestingly, the diglucosylated molecule, **EED′A′**, was the main glucosylation product. NMR analysis of this molecule revealed the presence of an unusual α-(1,6) linkage that might confer higher conformational flexibility to enable productive binding of **ED′A′** in subsites +1, +2, and+3. This could also explain to some extent why **ED′A′** appears to be a better acceptor than **D′A′**. Of note, similar observations were found in our recent work on the glucosylation of flavonoids.[42] These results also suggest a high plasticity of subsite +3 enabling accommodation of **A′**. Recent work on the engineering of the active site[43] also provides new directions to explore in order to better control the elongation of gluco-oligosaccharides and in particular, to avoid the formation of unwanted **EED′A′**.

CONCLUSIONS

Computer-aided design was for the first time applied to the engineering of several subsites of an α-transglycosylase. The difficulty of this work was to target a limited number of mutations allowing to keep a donor specificity for sucrose and tailor the acceptor specificity to render *Np*AS able to recognize a β-linked and partially protected disaccharide. Interestingly, *Np*AS has only been described up to now as being able to recognize α-linked disaccharides. The engineering approach included the use of RosettaDesign software, the selection of amino-acid sequences favoring inter-residue correlations predicted by computational design, and the experimental control of the desired combinations of mutations. The medium size library was easily screened to yield, after only a single screening run, one mutant displaying the expected substrate specificity, thus showing the efficacy of the approach combining both computer-aided design and construction of sequence libraries. Indeed, out of a library of 2.7×10^4 variants, the mutant F3 regioselectively glucosylates **D′A′**, what was not detected for the parental wild-type enzyme. A novel transglucosylase, whose substrate specificity was purposely tailored toward a non-natural acceptor, allyl 2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl-(1→2)-α-L-rhamnopyranoside (**D′A′**), to yield a potential intermediate in the chemoenzymatic synthesis of oligosaccharides representative of *S. flexneri* serotypes 1a, 1b, or 1d O-SPs. Obviously, although improvements are still necessary, these new findings open the way to alternatives to ongoing developments toward chemically synthesized oligosaccharides representing fragments from the O-SP of various *S. flexneri* type 1 strains.[51,52] However, the low hit rate emphasizes the extreme complexity of redesigning amylosucrase, given its double-step mechanism that involves the sequential binding of a donor (sucrose) and an acceptor substrate along the catalytic reaction, and the constrained active site topology which needs to be adapted to **D′A′** disaccharide which presents considerable structural differences with natural acceptors.

Indeed, the novel substrate specificity of the mutant was obtained upon introduction of 7 combined amino acid mutations in the first shell of the active site, enabling a drastic reshaping of the catalytic binding site to accommodate non-natural acceptor **D′A′** in a productive manner without perturbing too much its original specificity for donor sucrose. Not only these mutations, mainly introduced in flexible loops, enabled a change in substrate specificity but they also enhanced the stability of the enzyme in comparison to the parental wild-type *Np*AS. The catalytic efficiency of the mutant for trisaccharide production was still low compared to that observed for the natural donor sucrose. This is due to the difficulty of this design and can also be attributed to the limits of CPD approaches and experimental difficulties encountered in the construction of the library. Nonetheless, the combined approach used herein was efficient to identify in only one run of screening the evolutionary path toward a totally new and unnatural catalytic activity. Obviously, this opens now the route to further enzyme improvement via either computational methods, directed evolution or mixed strategies.

THEORETICAL AND EXPERIMENTAL PROCEDURES

Computational Modeling and Design. Detailed procedures are provided in Supporting Information and Methods. Three-dimensional model of *Np*AS in complex with the desired glucosylated product (**ED′A′**) bound into the enzyme active site was constructed starting from the crystallographic complex with maltoheptaose, a reaction product (PDB ID: 1MW0).[25] **ED′A′** was manually docked into the active site of the amylosucrase by superimposing the glucopyranosyl unit (**E**) of **ED′A′** onto the glucopyranosyl unit of the crystallographic maltoheptaose located at subsite −1. The molecular all-atom *ff99SB*,[44] *Glycam06*[45] and *gaff* force fields[46] were used for the proteins, carbohydrates, and other nonstandard groups, respectively. The molecular system was then subjected to several cycles of minimizations applying different levels of restraints using the *Sander* module of the Amber 9 software package.[47] Visual inspection of the modeled complex, as well as the crystallographic complex of the amylosucrase with the donor substrate, the sucrose (PDB: 1JGI), was carried out using PyMOL software (Schrödinger, Portland, OR, USA). Based on the analysis of these minimized complexes and our knowledge on amino acid residues involved in catalysis and/or specific recognition of sucrose, 23 positions (226−229, 289−300, 331, 332, 396−398, 434−439, 444−448) of the active site were selected to be sampled for mutations. Version 3.2 of the Rosetta software suite (Rosetta Commons, UW techTransfer, Seattle, WA, U.S.A.) was used to redesign the *Np*AS active site in order to enhance the interaction with **ED′A′**. A protocol was used that tests mutations and side-chain conformations at the 23 selected positions and samples variations in the translation and rotation of the docked **ED′A′**. All amino acid types, except Cysteine, were authorized at the 23 designable positions. Sixty residues surrounding the 23

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

designable residues were allowed to relax toward alternate side-chain rotamers. Enzyme backbone atoms, nondesigned or nonrepacked side-chains as well as the internal conformation of **ED′A′** were held fixed. The initial **ED′A′** conformation from which perturbations were made was taken from the minimized model of the **ED′A′**-enzyme complex. The design protocol based on the use of RosettaDesign and described in details in the Supporting Information and Methods was run 20 000 independent times. The output sequences were filtered based on different Rosetta scoring and filtering schemes. The 1515 unique sequences resulting from this filtering were used to guide library design. The library design approach developed was based on the following: (1) the determination from a multiple sequence alignment of 1515 designed sequences of the position-dependent amino acid residue frequency, a Correlated Mutation score at position pairs and the residue-pair frequency; (2) the calculation of the Shannon entropy using in house software, called Sequester, to measure the position-dependent amino acid residue variability in multiple sequence alignment of GH13_4 subfamily sequences.

From these computations, the amino acid types and the combinations of amino acid residues considered in the final designed libraries were selected as follows: (1) at each enzyme designed position, the wild-type amino acid residues with a frequency of occurrence ≥ 5% were maintained; (2) the wild-type amino acid residues with a frequency of occurrence < 5% or absent in the design results and with a Shannon entropy < 0.2% were introduced; (3) the other amino acid types were kept if only their position-dependent frequency was higher than 10%; (4) for correlated amino acid pairs, only the combination-pairs with a frequency of occurrence ≥ 5% were preserved, while all combinations were considered for noncorrelated amino-acid residues.

The mutant F3 of *Np*AS was constructed starting from the model of the wild-type enzyme docked with **ED′A′** in which mutations were introduced using the Biopolymer module of the InsightII software package (Accelrys, San Diego, CA, U.S.A.). The model was subsequently minimized following the same energy minimization procedure described above.

Chemical Synthesis of Disaccharide 1 (D′A′). General conditions and NMR spectra of all compounds are provided in Supporting Information.

*Allyl 3,4-O-(2′,3′-dimethoxybutane-2′,3′-diyl)-α-L-rhamnopyranoside (**4**).* To a solution of butane-2,3-dione (6.0 mL, 68.4 mmol) in MeOH (25 mL) were added trimethylorthoformate (22.4 mL, 204 mmol, 3.0 equiv) and three drops of concentrated sulfuric acid. The solution was refluxed for 20 h. After complete conversion of the starting dione, as attested by NMR analysis, solid $NaHCO_3$ (5.2 g) was added, and the suspension was concentrated to dryness under reduced pressure. The residue was dissolved in Et_2O (150 mL), washed with satd aq. $NaHCO_3$ (50 mL) and brine (50 mL), then dried over sodium sulfate. Volatiles were removed and the resulting material was distilled under reduce pressure to afford 2,2,3,3-tetramethoxybutane as a colorless oil (bp 35−37 °C, 4 mbar).

Distilled 2,2,3,3-tetramethoxybutane (4.15 g, 23.3 mmol, 1.2 equiv) and CSA (209 mg, 0.90 mmol, 0.05 equiv.) in anhydr. MeCN (3 mL) were added to a solution of the known rhamnoside[36] **3** (3.93 g, 19.2 mmol) in anhydr. MeCN (90 mL). The mixture was refluxed for 30 min. Et_3N was added and solvents were removed under reduced pressure. The crude oil was purified by column chromatography (Tol/EtOAc, 80:20 → 60:40) to give alcohol **4** (5.37 g, 88%) as a colorless oil. Alcohol **4** had R_f = 0.49 (cHex/EtOAc 3:2); $[\alpha]^{24}_D$ = −229 ° (*c* 1.0, $CHCl_3$); ^{1}H NMR ($CDCl_3$): δ 5.93−5.83 (m, 1H, C *H*=CH_2), 5.29−5.23 (m, J_{trans} = 17.2 Hz, 1H, CH=CH_aH_b), 5.19−5.16 (m, J_{cis} = 10.4 Hz, 1H, CH=CH_a H_b), 4.81 (d, $J_{1,2}$ = 1.2 Hz, 1H, H-1), 4.19−4.13 (m, 1H, -OCH_aH_{bAll}), 4.00−3.95 (m, 2H, -OCH_aH_{bAll}, H-3), 3.92 (dd, $J_{2,3}$ = 3.2 Hz, 1H, H-2), 3.82 (dq, $J_{4,5}$ = 9.8 Hz, 1H, H-5), 3.70 (pt, $J_{3,4}$ = 9.9 Hz, 1H, H-4), 3.26 (s, 3H, -OCH_3), 3.23 (s, 3H, -OCH_3), 2.42 (bs, 1H, OH), 1.31 (s, 3H, CH_3), 1.28 (s, 3H, CH_3), 1.26 (d, $J_{5,6}$ = 6.2 Hz, 3H, H-6). ^{13}C NMR ($CDCl_3$): δ 134.0 (*C*H=CH_2), 117.6 (CH=*C*H_2), 100.3 ($C_{quat.}$), 100.0 ($C_{quat.}$), 99.0 (C-1), 70.0 (C-2), 68.5 (C-4), 68.4 (C-3), 68.1 (-OCH_{2All}), 66.7 (C-5), 48.2 (-OCH_3), 47.8 (-OCH_3), 18.0 (CH_3), 17.8 (CH_3), 17.7 (C-6). HR-ESI-TOF-MS *m*/*z* 341.1542 (Calcd for $C_{15}H_{26}O_7Na$, 341.1576).

*Allyl (3,4,6-tri-O-acetyl-2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl)-(1→2)-3,4-O-(2′,3′-dimethoxybutane-2′,3′-diyl)-α-L-rhamnopyranoside (**6**).* TMSOTf (200 μL, 1.1 mmol, 0.1 equiv) was added to a solution of acceptor **4** (3.66 g, 11.5 mmol) and of the known donor[40] **5** (7.52 g, 12.6 mmol, 1.1 equiv) in anhydr. DCM (180 mL) containing 4Å MS (10 g), stirred at −15 °C under an Ar atmosphere. After 30 min, Et_3N was added, and the suspension was filtered over a pad of Celite. The filtrate was concentrated to dryness. Crystallization of the crude material from EtOAc/pentane gave disaccharide **6** (7.17 g, 9.54 mmol) as white crystals in a non optimized 83% yield. The fully protected compound **6** had R_f = 0.31 (cHex/EtOAc 3:2); $[\alpha]^{24}_D$ = −92.0° (*c* 1.0, $CHCl_3$); mp = 186 °C (Et_2O/pentane); ^{1}H NMR ($CDCl_3$): δ 6.83 (d, $J_{2,NH}$ = 7.2 Hz, 1H, NH), 5.91−5.82 (m, 1H, *CH*=CH_2), 5.31−5.22 (m, J_{trans} = 17.2 Hz, 2H, CH=CH_2, H-3$_{D'}$), 5.17−5.13 (m, J_{cis} = 10.4 Hz, 1H, CH=CH_2), 5.08 (pt, J = 9.5 Hz, 1H, H-4$_{D'}$), 4.98 (d, $J_{1,2}$ = 8.2 Hz, $^1J_{CH}$ = 164.4 Hz, 1H, H-1$_{D'}$), 4.86 (d, $J_{1,2}$ = 1.2 Hz, 1H, H-1$_{A'}$), 4.24 (dd, $J_{6a,6b}$ = 12.1 Hz, $J_{5,6a}$ = 5.1 Hz, 1H, H-6a$_{D'}$), 4.18−4.10 (m, 2H, H-6b$_{D'}$, H_{All}), 4.08−4.00 (m, 2H, H-2$_{D'}$, H-3$_{A'}$), 3.98−3.92 (m, 2H, H_{All}, H-2$_{A'}$), 3.76−3.67 (m, 2H, H-5$_{A'}$, H-5$_{D'}$), 3.56 (pt, J = 9.9 Hz, 1H, H-4$_{A'}$), 3.23 (s, 3H, -OCH_3), 3.17 (s, 3H, -OCH_3), 2.08 (s, 3H, CH_{3Ac}), 2.02 (s, 3H, CH_{3Ac}), 2.01 (s, 3H, CH_{3Ac}), 1.25 (s, 3H, CH_3), 1.24 (s, 3H, CH_3), 1.21 (d, $J_{5,6}$ = 6.2 Hz, 3H, H-6$_{A'}$); ^{13}C NMR ($CDCl_3$): δ 170.8, 170.7, 169.4 (3C, CO_{Ac}), 161.9 (NHCO), 134.0 (*C*H=CH_2), 117.3 (CH=*C*H_2), 100.5 ($C_{quat.}$), 100.2 ($C_{quat.}$), 99.7 (C-1$_{D'}$, $^1J_{C,H}$ = 164.1 Hz), 99.0 (C-1$_{A'}$, $^1J_{C,H}$ = 173.8 Hz), 92.4 (CCl_3), 76.1 (C-2$_{A'}$), 72.2 (2C, C-3$_{D'}$, C-5$_{D'}$), 68.7 (2C, C-4$_{A'}$, C-4$_{D'}$), 68.1 (2C, C-3$_{A'}$, C_{All}), 66.8 (C-5$_{A'}$), 62.2 (C-6$_{D'}$), 56.0 (C-2$_{D'}$), 48.0 (-OCH_3), 47.8 (-OCH_3), 20.8 (CH_{3Ac}), 20.7 (2C, CH_{3Ac}), 18.0 (2C, CH_3), 16.7 (C-6$_{A'}$). HR-ESI-TOF-MS *m*/*z* 772.1539 (Calcd for $C_{29}H_{42}Cl_3NO_{15}Na$, 772.1518).

*Allyl (3,4,6-tri-O-acetyl-2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl)-(1→2)-α-L-rhamnopyranoside (**7**).* Disaccharide **6** (1.94 g, 2.58 mmol) was dissolved in 90% aq. TFA (50 mL), and the solution was stirred for 50 min at rt, then repeatedly coevaporated with toluene. The residue was purified by column chromatography (cHex/EtOAc 50:50 → 40:60) to give diol **7** (1.41 g, 2.21 mmol, 86%) as a white solid. Disaccharide **7** had R_f = 0.38 (cHex/EtOAc 2:3); $[\alpha]^{24}_D$ = −24.4° (*c* 1.0, $CHCl_3$); ^{1}H NMR ($CDCl_3$): δ 7.40 (d, $J_{2,NH}$ = 8.4 Hz, 1H, NH), 5.92−5.83 (m, 1H, *CH*=CH_2), 5.29−5.24 (m, J_{trans} = 17.2 Hz, 2H, CH=CH_2, H-3$_{D'}$), 5.19−5.16 (m, J_{cis} = 10.4 Hz, 1H, CH=CH_2), 5.09 (pt, J = 9.8 Hz, 1H, H-4$_{D'}$), 4.94 (d, $J_{1,2}$ = 8.4 Hz, 1H, H-1$_{D'}$), 4.89 (d, $J_{1,2}$ = 1.3 Hz, 1H, H-1$_{A'}$), 4.24 (dd, $J_{6a,6b}$ = 12.4 Hz, $J_{5,6a}$ = 5.0 Hz, 1H, H-6a$_{D'}$), 4.17−4.06 (m, 3H, H-6b$_{D'}$, H_{All}, H-2$_{D'}$), 3.97−3.93 (m, 2H, H_{All}, H-2$_{A'}$), 3.85 (dd, $J_{3,4}$ = 9.5 Hz, $J_{2,3}$ = 3.3 Hz, 1H, H-3$_{A'}$), 3.76−3.72 (m, 1H, H-5$_{D'}$),

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

3.56−3.36 (dq, $J_{4,5}$ = 9.4 Hz, 1H, H-$5_{A'}$), 3.41 (pt, J = 9.5 Hz, 1H, H-$4_{A'}$), 2.11 (s, 3H, CH_{3Ac}), 2.06 (s, 3H, CH_{3Ac}), 2.05 (s, 3H, CH_{3Ac}), 1.30 (d, $J_{5,6}$ = 6.2 Hz, 3H, H-$6_{A'}$); ^{13}C NMR ($CDCl_3$): δ 171.0, 170.8, 169.5 (3C, CO_{Ac}), 162.4 (NHCO), 133.9 (*C*H═CH_2), 117.4 (CH═CH_2), 102.0 (C-$1_{D'}$, $^1J_{C,H}$ = 163.4 Hz), 98.3 (C-$1_{A'}$, $^1J_{C,H}$ = 173.2 Hz), 92.5 (CCl_3), 79.2 (C-$2_{A'}$), 73.7 (C-$4_{A'}$), 72.3 (2C, C-$3_{D'}$, C-$5_{D'}$), 71.8 (C-$3_{A'}$), 68.5 (C-$4_{D'}$), 68.2 (C-$5_{A'}$), 68.1 (C_{All}), 62.2 (C-$6_{D'}$), 56.2 (C-$2_{D'}$), 20.9, 20.7, 20.6 (3C, CH_{3Ac}), 17.5 (C-$6_{A'}$). HR-ESI-TOF-MS m/z 658.0876 (Calcd for $C_{23}H_{32}Cl_3NO_{13}Na$, 658.0837).

*Allyl (2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl)-(1→2)-α-L-rhamnopyranoside (**1** or **D′A′**).* To a solution of diol **7** (1.27 g, 2.0 mmol) in anhydr. MeOH (30 mL) stirred under an Ar atmosphere, was added MeONa (25 wt % in MeOH, 275 μL, 1.20 mmol, 0.6 equiv). The reaction mixture was stirred at rt for 2 h, then Dowex 50Wx8-200 resin (H+ form) was added. The suspension was filtered over a pad of Celite. The solvent was removed under vacuum and the residue was purified by column chromatography (DCM/MeOH, 90:10 → 80:20) to give pentaol **1** (937 mg, 1.83 mmol, 92%) as a white solid. Disaccharide **1** had R_f = 0.11 (DCM/MeOH 9:1); $[\alpha]^{24}_D$ = −29.3° (*c* 1.0, water); ^{1}H NMR (D_2O): δ 6.08−5.98 (m, 1H, *CH*═CH_2), 5.45−5.39 (m, J_{trans} = 17.4 Hz, 1H, CH═*CH*$_2$), 5.37−5.33 (m, J_{cis} = 10.4 Hz, 1H, CH═*CH*$_2$), 5.06 (d, $J_{1,2}$ = 1.7 Hz, 1H, H-$1_{A'}$), 4.98 (d, $J_{1,2}$ = 8.1 Hz, 1H, H-$1_{D'}$), 4.30−4.24 (m, 1H, H_{All}), 4.16−4.11 (m, 2H, H_{All}, H-$2_{A'}$), 4.00 (bd, $J_{6a,6b}$ = 12.5 Hz, 1H, H-$6a_{D'}$), 3.91 (dd, $J_{3,4}$ = 9.6 Hz, $J_{2,3}$ = 3.1 Hz, 1H, H-$3_{A'}$), 3.88−3.80 (m, 3H, H-$6b_{D'}$, H-$2_{D'}$, H-$3_{D'}$), 3.80−3.72 (m, 1H, H-$5_{A'}$), 3.57−3.55 (m, 2H, H-$4_{D'}$, H-$5_{D'}$), 3.50 (pt, J = 9.6 Hz, 1H, H-$4_{A'}$), 1.33 (d, $J_{5,6}$ = 6.3 Hz, 1H, H-$6_{A'}$). ^{13}C NMR (D_2O): δ 164.9 (NHCO), 133.3 (*C*H═CH_2), 118.5 (CH═*CH*$_2$), 101.4 (C-$1_{D'}$, $^1J_{C,H}$ = 164.4 Hz), 97.9 (C-$1_{A'}$, $^1J_{C,H}$ = 172.9 Hz), 91.6 (CCl_3), 78.0 (C-$2_{A'}$), 75.9 (C-$4_{D'}$), 73.2 (C-$3_{D'}$), 72.4 (C-$4_{A'}$), 70.1 (C-$3_{A'}$), 70.0 (C-$5_{D'}$), 68.8 (C-$5_{A'}$), 68.3 (C_{All}), 60.6 (C-$6_{D'}$), 57.8 (C-$2_{D'}$), 16.6 (C-$6_{A'}$); HR-ESI-TOF-MS m/z 532.0504 (Calcd for $C_{17}H_{26}Cl_3NO_{10}Na$, 532.0520).

Bacterial Strains, Plasmids, and Chemicals. Details are provided in Supporting Information and Methods.

Library Construction. The gene library was constructed in five PCR steps using Phusion^R polymerase (New England Biolabs, Ipswich, MA, U.S.A.) (Figure S2 of the Supporting Information). In the first step, the six DNA wild-type fragments surrounding the five regions containing designed mutations were amplified (Figure S2 of the Supporting Information).

Amplified products were purified by a gel extraction kit. In a second step, PCR reactions were divided in 10 pools (Figure S2 of the Supporting Information). In this way, in pool 1, the fragment F1 (10 ng) was mixed with 226−229_rev (0.5 μM), the reverse primer containing mutated codons for 226, 227, 228, and 229 positions. In pool 2, the fragment F2 was mixed with 226−229_for, the forward primer containing mutated codons for 226, 227, 228, and 229 positions. In pool 3, the fragment F2 was mixed with 289−300_rev, the reverse primer containing mutated codons for 289, 290, and 300 positions, and so on for the 10 pools. PCR was conducted in a total volume of 20 μL in the following conditions: one step at 98 °C for 30 s, followed by 20 cycles at 98 °C-10 s, 52 °C-30 s (pool 1 to 10), 72 °C-20 s. DNA fragments and primers used in each pool contained overlapping regions of at least 20 nucleotides. The primers containing degenerated codons were specific to each library (Table S1 of the Supporting Information). In a third step, amplicons obtained from step 2 were mixed two by two to perform 1 cycle at 98 °C for 30 s and 10 cycles at 98 °C for 10 s and 72 °C for 40 s. In a fourth step, 10 μL of each mixture coming from the third step were mixed together to perform 1 cycle at 98 °C for 30 s and 10 cycles at 98 °C for 10 s and 72 °C for 1 min. No Phusion polymerase was added in steps 3 and 4. In a fifth step, the gene library obtained was finally amplified by adding outer primers (see Supporting Information).

Amplified products were purified by a gel extraction kit. Amylosucrase libraries were cloned into pGEX-6P-3 (GE Healthcare Biosciences) using *Eco*RI HF and NotI HF restriction sites. Ligation products were transformed in *E.coli* TOP 10 electrocompetent cells (Invitrogen, Carlsbad, U.S.A.). Freshly transformed cells were plated on lysogeny broth (LB) agar supplemented with 100 μg/mL ampicillin for 24 h at 37 °C.

Library Screening on Sucrose in Solid Media. A pH-based high-throughput screening assay earlier described [35] was used to isolate sucrose active clones (see Supporting Information and Methods provided in Supporting Information).

Expression of Active Clones in Microplate Format. Active mutants were replicated to inoculate a 96-well microplates containing LB medium (200 μL), ampicillin (100 μg/mL^{-1}), and glycerol (12% w/v) and incubated at 30 °C for 24h under agitation at 450 rpm (orbital diameter of 3 mm). Plates were then duplicated into 96-deep-well plates containing in each well 1 mL of autoinducing media ZYM-5052 [48] supplemented with ampicillin (100 μg/mL) for culture and gene expression. These plates were incubated 24 h at 30 °C with agitation at 700 rpm. To evaluate the growth homogeneity, absorbance was measured at 600 nm for each culture diluted 10 times in water (Sunrise spectrophotometer, Tecan, Männedorf, Switzerland). Plates were then centrifuged (20 min, 3000*g*, 4 °C), supernatants were removed, and the cell pellets were resuspended in 200 μL of 2× PBS buffer containing lysozyme (0.5 mg/mL) and DNase and frozen at −80 °C overnight. After thawing at room temperature, microplates were centrifuged (20 min, 3000*g*, 4 °C); soluble fractions were then transferred in new microplates for reaction assays.

Reaction assays in the presence of sucrose were carried out as previously described.[19,20] All 55 positive hits were sequenced on the entire gene by GATC Biotech GA (Constance, Germany).

Reaction Assays with Sucrose and D′A′ Acceptor in Microplate Format. Reactions were carried out in microplates with 100 μL of soluble fractions and 50 μL of sucrose solution (73 mM final concentration) and 50 μL of **D′A′** acceptor solution (73 mM final concentration). The enzymatic reaction was incubated at 30 °C for 24 h under agitation (450 rpm). The reactions were then stopped by heating at 95 °C for 5 min, centrifugated (20 min at 3700 rpm), filtered and analyzed by HPLC using conditions described hereafter.

Carbohydrate Analysis. Sucrose consumption was evaluated by ion exchange chromatography at 65 °C using an Aminex HP87K column (300 × 7.8 mm, Biorad) with ultrapure water at 0.6 mL/min as eluent. The detection was performed using a shodex RI 101 series refractometer. **D′A′** consumption and **ED′A′** and **EED′A′** production were identified using a reversed C18 phase column (Venusil AQ 50 mm) with MeCN/water (17:83, v/v) as eluent. The detection was performed using a UV detector (Dionex, UVD 340) at 210 and 220 nm. The glucosylated molecule was identified by LC-MS with the above conditions. Using a MSQ Plus Mass Spectrometer (Dionex - Thermoscientific equipment), molecules were ionised by electrospray ionization technique with a source at 450 °C. Separation was achieved with a quadripole and detected using the negative mode.

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

Production and Purification of Isolated Mutants. *E. coli* TOP10 electrocompetent cells containing pGST-AS mutated gene were produced as previously described[49] but with a preculture step at 30 °C overnight instead of 37 °C and then cultures at 23 °C. After sonication, the crude extract was centrifugated (10 000 rpm, 20 min, 4 °C) and the supernatant was harvested for the purification step. *Np*AS mutants were purified on an affinity chromatography column via affinity of the GST tag to the Glutathion Sepharose 4B support (Dutscher, Brumath, France) into Tris buffer (50 mM Tris, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.0). GST Tag was removed by PreScission protease (Dutscher, Brumath, France). Protein concentration was estimated using Nanodrop ND-1000 spectrophotometer, and specific activity was determined at 30 °C, in 50 mM Tris-HCl buffer pH 7.0, using 50 g/L sucrose during 20 min by dinitrosalicylic acid assay. One unit of amylosucrase activity corresponds to the amount of enzyme that catalyzes the release of 1 μmol of reducing sugars per minute in the assay conditions.

Characterization of Product Profiles of the Active Mutant on D′A′. Reactions were performed at 30 °C during 24 h with 146 mM sucrose alone or supplemented with 146 mM acceptor molecule. Wild-type enzyme and mutants were employed at 1 U/mL, and reaction mixes were stopped by heating at 95 °C during 5 min. Product profiles were analyzed by HPAEC-PAD using a Dionex Carbo-Pack PA100 column at 30 °C. Soluble fraction was analyzed with a gradient of sodium acetate (from 6 to 300 mM in 28 min) in 150 mM NaOH and the crude extract was solubilized in 1 M aqueous KOH. Detection was performed using a Dionex ED40 module with a gold working electrode and Ag/AgCl reference electrode.

Kinetic Parameters Determination of Active Mutant. Enzyme kinetics were carried out with pure enzymes (0.2 and 0.1 U/ml) at 30 °C, under agitation. Catalytic efficiency of wild-type *Np*AS and mutants was determined with sucrose alone from 0 to 600 mM and with both sucrose (146 mM) and acceptor molecule (from 5 mM to 500 mM) as variable substrates. Catalytic efficiencies were determined by measuring the release of fructose monitored for sucrose alone ($k_{cat}/K_{m,\,suc}$), and by measuring the rate of formation ($k_{cat}/K_{m,suc}$; $k_{cat}/K_{m,D'A'}$) of the monoglucosylated disaccharide. After heating at 95 °C, and centrifugation, aliquots were analyzed by HPLC RI with a Biorad HPX-87K column.

Semipreparative Synthesis and Identification of Allyl α-D-glucopyranosyl-(1→6)-α-D-glucopyranosyl-(1→4)-2-deoxy-2-trichloroacetamido-β-D-glucopyranosyl-(1→2)-α-L-rhamnopyranoside (E′ED′A′). A mixture of 146 mM **D′A′**, 292 mM sucrose, and 1 U/mL mutant F3 (2 mL) was gently stirred at 30 °C. After 24 h, the reaction mix was heated at 95 °C for 5 min, and then centrifugated (6000 rpm, 4 °C, 20 min). The glucosylated products were separated by reverse-phase semipreparative chromatography using a C18 column (Bishoff Chromatography, 250 mm × 8 mm), eluting with 17% acetonitrile in pure water at a constant flow rate of 2 mL/min at 30 °C. The major product was lyophilized repeatedly to give tetrasaccharide **EED′A′**. This product had ^{1}H NMR (D_2O): δ 5.99−5.89 (m, 1H, C*H*═CH$_2$), 5.43 (d, $J_{1,2}$ = 3.8 Hz, 1H, H-1$_E$), 5.35−5.31 (m, J_{trans} = 17.3 Hz, 1H, CH═C*H*$_2$), 5.27−5.25 (m, J_{cis} = 10.5 Hz, 1H, CH═C*H*$_2$), 4.98 (bs, 1H, H-1$_{A'}$), 4.94 (d, $J_{1,2}$ = 3.6 Hz, 1H, H-1$_{E'}$), 4.90 (d, $J_{1,2}$ = 8.4 Hz, 1H, H-1$_{D'}$), 4.19 (m, 1H, H$_{All}$), 4.07−4.01 (m, 3H, H-2$_{A'}$, H-3$_{D'}$, H$_{All}$), 3.99−3.91 (m, 2H, H-6a$_{E'}$, H-6a$_{D'}$), 5.88 (m, 1H, H-5$_{E'}$), 3.84−3.77 (m, 3H, H-6a$_{D'}$, H-3$_{A'}$, H-5$_E$), 3.75−3.68 (m, 5H, H-6a$_E$, H-3$_E$, H-6b$_E$, H-6b$_{E'}$, H-4$_{D'}$), 3.66−3.63 (m, 2H, H-3$_{E'}$, H-5$_{A'}$), 3.60−3.55 (m, 2H, H-2$_{E'}$, H-5$_{A'}$), 3.51 (dd$_{po}$, 1H, $J_{2,3}$ = 10.4 Hz, 1H, H-2$_E$), 3.48 (pt$_{po}$, 1H, $J_{3,4}$ = $J_{4,5}$ = 9.3 Hz, H-4$_{E'}$), 3.41 (pt$_{po}$, 1H, $J_{3,4}$ = $J_{4,5}$ = 9.2 Hz, H-4$_E$), 3.40 (pt$_{po}$, 1H, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4$_{A'}$), 1.23 (d, $J_{5,6}$ = 6.5 Hz, 1H, H-6$_{A'}$); ^{13}C NMR (D_2O): δ 167.5 (NHCO), 136.0 (*C*H═CH$_2$), 121.1 (CH═*C*H$_2$), 104.0 (C-1$_{D'}$, $^1J_{C,H}$ = 164.1 Hz), 102.1 (C-1$_{E'}$, $^1J_{C,H}$ = 172.7 Hz), 100.7 (C-1$_E$, $^1J_{C,H}$ = 171.8 Hz), 100.5 (C-1$_{A'}$, $^1J_{C,H}$ = 173.7 Hz), 94.2 (*C*Cl$_3$), 80.7 (C-2$_{A'}$), 79.3 (C-3$_E$), 77.2 (C-5$_{D'}$), 76.3 (C-3$_{D'}$), 75.8 (C-3$_{A'}$), 75.7 (C-4$_{D'}$), 75.0 (C-4$_{A'}$), 74.5 (C-3$_E$), 74.1 (2C, C-2$_{E'}$, C-2$_E$), 74.0 (C-5$_{E'}$), 72.7 (C-5$_E$), 72.2 (C-4$_{E'}$), 72.0 (C-4$_E$), 71.4 (C-5$_{A'}$), 70.9 (C$_{All}$), 68.5 (C-6$_{E'}$), 63.3 (C-6$_{D'}$), 63.1 (C-6$_E$), 60.4 (C-2$_{D'}$), 19.2 (C-6$_{A'}$). *m/z* (ES-) for **EED′A′**, *m/z* 832 [M - H$^+$].

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The following file is available free of charge on the ACS Publications website at DOI: 10.1021/cs501288r.

Supplementary tables and figures showing details of library design, construction, screening and analysis; general methods for chemical synthesis; ^{1}H and ^{13}C NMR spectra for all novel compounds (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: isabelle.andre@insa-toulouse.fr.

*E-mail: remaud@insa-toulouse.fr.

Present Address

□(For S.S.) Oxeltis, Cap Gamma − 1682, rue de la Valsière, CS 27384, 34189 Montpellier, France

Author Contributions

#A.V., E.C., and S.B. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to C. Topham for providing his help with Sequester analysis. They also thank G. Cioci and S. Tranier in using the PICT facility for protein biophysical characterization, N. Monties for her help in the product analysis and purification by HPAEC-PAD and LC-MS as well as S. Bozonnet and S. Pizzut-Serin in using the ICEO high-throughput facility, devoted to the engineering and screening of new and original enzymes, of the Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (Toulouse, France). The authors wish to thank C. Guerreiro (Unité de Chimie des Biomolécules, Institut Pasteur, Paris, France) for her technical contribution to the synthetic aspects of this work, and F. Bonhomme (CNRS, UMR 3523, Institut Pasteur, Paris, France) for providing the HRMS data. This work was granted access to the HPC resources of the Computing Center of Region Midi-Pyrénées (CALMIP, Toulouse, France). This work was supported by the French National Research Agency (ANR Project GLUCODESIGN, ANR-08-PCVI-002-02) and by the European Commission Seventh Framework Program (FP7/2007−2013) under Grant agreement No. 261472-STOPENTERICS

■ REFERENCES

(1) Turner, N. J.; Truppo, M. D. *Curr. Opin. Chem. Biol.* **2013**, *17*, 212−214.

(2) Turner, N. J.; O'Reilly, E. *Nat. Chem. Biol.* **2013**, *9*, 285−288.

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198

(3) Davids, T.; Schmidt, M.; Böttcher, D.; Bornscheuer, U. T. *Curr. Opin. Chem. Biol.* **2013**, *17*, 215−220.
(4) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. *Nature* **2012**, *485*, 185−194.
(5) Höhne, M.; Schätzle, S.; Jochens, H.; Robins, K.; Bornscheuer, U. T. *Nat. Chem. Biol.* **2010**, *6*, 807−813.
(6) Kiss, G.; Çelebi-Ölçüm, N.; Moretti, R.; Baker, D.; Houk, K. N. *Angew. Chem., Int. Ed.* **2013**, *52*, 5700−5725.
(7) Hilvert, D. *Annu. Rev. Biochem.* **2013**, *82*, 447−470.
(8) Kries, H.; Blomberg, R.; Hilvert, D. *Curr. Opin. Chem. Biol.* **2013**, *17*, 221−228.
(9) Wijma, H. J.; Janssen, D. B. *FEBS J.* **2013**, *280*, 2948−2960.
(10) Voigt, C. A.; Mayo, S. L.; Arnold, F. H.; Wang, Z.-G. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3778−3783.
(11) Voigt, C. A.; Mayo, S. L.; Arnold, F. H.; Wang, Z. *J. Cell. Biochem. Suppl.* **2002**, *37*, 58−63.
(12) Wijma, H. J.; Floor, R. J.; Jekel, P. A.; Baker, D.; Marrink, S. J.; Janssen, D. B. *Protein Eng. Des. Sel.* **2014**, *27*, 49−58.
(13) Goldsmith, M.; Tawfik, D. S. Enzyme Engineering by Targeted Libraries. In *Methods in Enzymology*; Keating, A. E., Ed.; Academic Press: New York, 2013; Vol. *523*, pp 257−283.
(14) Allen, B. D.; Nisthal, A.; Mayo, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19838−19843.
(15) Lippow, S. M.; Moon, T. S.; Basu, S.; Yoon, S.; Li, X.; Chapman, B. A. *Chem. Biol.* **2010**, *17*, 1306−1315.
(16) Chaparro-riggers, J. F.; Polizzi, K. M.; Bommarius, A. S. *Biotechnol. J.* **2007**, *2*, 180−191.
(17) Floor, R. J.; Wijma, H. J.; Colpa, D. I.; Ramos-Silva, A.; Jekel, P. A.; Szymański, W.; Feringa, B. L.; Marrink, S. J.; Janssen, D. B. *ChemBioChem* **2014**, *15*, 1660−1672.
(18) André, I.; Potocki-Véronèse, G.; Barbe, S.; Moulis, C.; Remaud-Siméon, M. *Curr. Opin. Chem. Biol.* **2014**, *19*, 17−24.
(19) Champion, E.; André, I.; Moulis, C.; Boutet, J.; Descroix, K.; Morel, S.; Monsan, P.; Mulard, L. A.; Remaud-Siméon, M. *J. Am. Chem. Soc.* **2009**, *131*, 7379−7389.
(20) Champion, E.; Guérin, F.; Moulis, C.; Barbe, S.; Tran, T. H.; Morel, S.; Descroix, K.; Monsan, P.; Mourey, L.; Mulard, L. A.; Tranier, S.; Remaud-Siméon, M.; André, I. *J. Am. Chem. Soc.* **2012**, *134*, 18677−18688.
(21) Perepelov, A. V.; Shekht, M. E.; Liu, B.; Shevelev, S. D.; Ledov, V. A.; Senchenkova, S. N.; L'vov, V. L.; Shashkov, A. S.; Feng, L.; Aparin, P. G.; Wang, L.; Knirel, Y. A. *FEMS Immunol. Med. Microbiol.* **2012**, *66*, 201−210.
(22) Shashkov, A. S.; Senchenkova, S. N.; Sun, Q.; Lan, R.; Wang, J.; Perepelov, A. V.; Knirel, Y. A.; Xu, J. *Carbohydr. Res.* **2013**, *24*, 93−96.
(23) Skov, L. K.; Mirza, O.; Henriksen, A.; Potocki de Montalk, G.; Remaud-Siméon, M.; Sarçabal, P.; Willemot, R.; Monsan, P.; Gajhede, M. *J. Biol. Chem.* **2001**, *276*, 25273−25278.
(24) Skov, L. K.; Mirza, O.; Gajhede, M.; Feller, G.; D'Amico, S.; André, G.; Potocki-Veronese, G.; van der Veen, B.; Monsan, P.; Remaud-Simeon, M. *J. Biol. Chem.* **2004**, *279*, 726−734.
(25) Skov, L. K.; Mirza, O.; Sprogøe, D.; Dar, I.; Remaud-Simeon, M.; Albenne, C.; Monsan, P.; Gajhede, M. *J. Biol. Chem.* **2002**, *277*, 47741−47747.
(26) Hanson, S.; Best, M.; Bryan, M. C.; Wong, C-H. *Trends Biochem. Sci.* **2004**, *29*, 656−663.
(27) Wang, Z.; Chinoy, Z.; Ambre, S.; Peng, W; McBride, R.; de Vries, R. P.; Glushka, J.; Paulson, J. C.; Boons, G. J. *Science* **2013**, *341*, 379−383.
(28) Armstrong, Z.; Withers, S. G. *Biopolymers* **2009**, *99*, 666−674.
(29) Muthana, S.; Cao, H.; Chen, X. *Curr. Opin. Chem. Biol.* **2009**, *13*, 573−581.
(30) Lucas, R.; Hamza, D.; Lubineau, A.; Bonnaffé, D. *Eur. J. Org. Chem.* **2004**, *10*, 2107−2117.
(31) Boutet, J.; Guerreiro, C.; Mulard, L. A. *J. Org. Chem.* **2009**, *74*, 2651−2670.
(32) Chassagne, P.; Raibaut, L.; Guerreiro, C.; Mulard, L. A. *Tetrahedron* **2013**, *69*, 10337−10350.
(33) Reetz, M. T.; Kahakeaw, D.; Lohmer, R. *ChemBioChem* **2008**, *9*, 1797−1804.
(34) Stam, M. R.; Danchin, E. G. J.; Rancurel, C.; Coutinho, P. M.; Henrissat, B. *Protein Eng. Des. Sel.* **2006**, *19*, 555−562.
(35) Champion, E.; Moulis, C.; Morel, S.; Mulard, L. A.; Monsan, P.; Remaud-Siméon, M.; André, I. *ChemCatChem* **2010**, *2*, 969−975.
(36) Pinto, B. M.; Morisette, D. G.; B., D. R. *J. Chem. Soc., Perkin Trans.* **1987**, *1*, 9−14.
(37) Ley, S. V.; Priepke, H. W. M.; Warriner, S. L. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 2290−2292.
(38) Ley, S. V.; Polara, A. *J. Org. Chem.* **2007**, *72*, 5943−5959.
(39) Montchamp, J.-L.; Tian, F.; Hart, M. E.; Frost, J. W. *J. Org. Chem.* **1996**, *61*, 3897−3899.
(40) Boutet, J.; Kim, T. H.; Guerreiro, C.; Mulard, L. A. *Tetrahedron Lett.* **2008**, *49*, 5339−5342.
(41) Mirza, O.; Skov, L. K.; Remaud-Simeon, M.; Potocki de Montalk, G. P.; Albenne, C.; Monsan, P.; Gajhede, M. *Biochemistry* **2001**, *40*, 9032−9039.
(42) Malbert, Y.; Pizzut-Serin, S.; Massou, S.; Cambon, E.; Laguerre, S.; Monsan, P.; Lefoulon, F.; Morel, S.; André, I.; Remaud-Simeon, M. *ChemCatChem* **2014**, *16*, 2282−2291.
(43) Cambon, E.; Barbe, S.; Pizzut-Serin, S.; Remaud-Simeon, M.; André, I. *Biotechnol. Bioeng.* **2014**, *111*, 1719−1728.
(44) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins Struct. Funct. Genet* **2006**, *65*, 712−725.
(45) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622−655.
(46) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(47) Case, D. A.; Darde, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollmann, P. A. *AMBER 9*; University of California: San Francisco, CA, 2006
(48) Studier, F. W. *Protein Expr. Purif.* **2005**, *41*, 207−234.
(49) Potocki de Montalk, G.; Remaud-Siméon, M.; Willemot, R. M.; Planchot, V.; Monsan, P. *J. Bacteriol.* **1999**, *181*, 375−381.
(50) Lemanski, G.; Ziegler, T. *Eur. J. Org. Chem.* **2006**, 2618−2630.
(51) Dhara, D.; Kar, R. K.; Bhunia, A.; Misra, A. K. *Eur. J. Org. Chem.* **2014**, 4577−4584.
(52) Hargreaves, J.; Le Guen, Y.; Guerreiro, C.; Descroix, K.; Mulard, L. A. *Org. Biomol. Chem.* **2014**, *12*, 7728−7749.

DOI: 10.1021/cs501288r
ACS Catal. 2015, 5, 1186−1198